BCS THE CHARTERED INSTITUTE FOR IT

BCS HIGHER EDUCATION QUALIFICATIONS
BCS Level 6 Professional Graduate Diploma in IT

ADVANCED DATABASE MANAGEMENT SYSTEMS

Friday 24th March 2017 - Afternoon
Answer **any** THREE questions out of FIVE. All questions carry equal marks.
Time: THREE hours

Answer any Section A questions you attempt in Answer Book A
Answer any Section B questions you attempt in Answer Book B

The marks given in brackets are **indicative** of the weight given to each part of the question.

Calculators are **NOT** allowed in this examination.

Section A
Answer Section A questions in Answer Book A

A1 The definition of a true distributed database system is that it consists of a collection of geographically remote sites of physical databases, connected together via some kind of communication network, in which a user at any site can access data anywhere in the network exactly as if the data were all stored in one single database at the user's own site.

a) Explain the problems associated with achieving a true distributed database in practice. **(12 marks)**

b) Study the following scenario and then attempt the question parts that follow:

A small Bank has two branches, one located in York and one located in Leeds. Currently the banks centralised database is managed at its HQ in London, where it keeps data about its customer accounts. Other uses include collecting data for the generation of reports and to monitor the services that customers use. Applications have been installed at the two branches to access the centralised database via a fast communication network for accessing the data they need. There is also a communication link between the two branches, which is currently used only when one of the main links to the London headquarters fail.
For simplicity the centralised database system holds a single Customer table (see figure A1 below), where data about customer accounts are kept. The columns of the Customer table are, the account number, the customer's name, the branch where the account is kept and its current balance. The bank has been asked to split its branches so that they can support customers when there is failure of the centralised database or the communication system is slow or breaks down.

Figure 1 Examples of a customer relation held at HQ

Account_Number	Customer_Name	Branch	Balance
20012	Brown	York	1000.11
20454	Patel	Leeds	-250.89
61344	Smith	Leeds	58.99
n75630	Gray	York	2956.93
30435	Green	Leeds	-33.62
96567	Richards	York	45.76
01232	Akhtar	York	43.92

Describe **three** ways to distribute data across the three sites. Comment on the pros and cons of each option. **(13 marks)**

A2 This question considers advanced use of SQL in different programming contexts.

a) The following programming techniques can be thought of as extending the range and functionality of SQL and perform tasks that SQL cannot perform on its own. Describe with the aid of examples how ***each*** programming technique extends the range and functionality of SQL.

i) Stored procedures **(5 marks)**
ii) Triggers **(5 marks)**
iii) Embedding SQL in a programming language such as Java/PHP **(5 marks)**

b) Another programming technique is to use a native Object to Relational (OR) mapping language (eg LINQ) to access a relational database. Outline the principles behind an OR mapping language and outline the benefits of using this technique compared with a conventional programming technique such as embedding SQL in Java (part b iii)) above. Use examples to assist your answer.

(10 marks)

A3

a) Consider the following tables:

```
Film (filmNbr, title, year)
Director (directID, name)
Directing (directID, filmNbr)
```

And the following query:

```
SELECT Film.title
FROM  Film, Director, Directing
WHERE Film.filmNbr = Directing.filmNbr
AND Director.directID = Directing.directID
AND Director.name = 'Lucas'
AND Film.year = 2015;
```

(i) Draw a query tree that corresponds to the most efficient way of processing this query.

(10 marks)

(ii) Assume there is a B-Tree index on the column "title" of the table "Film". For each of the following queries, explain how this index could be used when executing each query:

```
SELECT * FROM Film WHERE title = 'Up';
SELECT * FROM Film ORDER BY title;
SELECT COUNT(title) FROM Film;
```

(6 marks)

b) The security of a database should be built-into the database development lifecycle. For each of the following database development stages, describe how security can be considered:

(i) Database Planning Stage **(2 marks)**
(ii) Database Requirements Stage **(2 marks)**
(iii) Database Design Stage **(2 marks)**

c) A medical surgery keeps patients' records in a database server located in a dedicated room. These records can be accessed remotely by the staff working at the surgery (doctors and office staff).

(i) Describe an example of physical security control for protecting this database. **(1 mark)**

(ii) Describe one way of preventing the office staff from viewing sensitive medical data about patients that should be viewed only by doctors. **(2 marks)**

Section B
Answer Section B questions in Answer Book B

B4

Using your own simple examples and suitable diagrams, explain the following transaction-processing concepts.

a) Guaranteed data consistency via ACID principles. **(5 marks)**

b) Eventual data consistency via BASE principles. **(5 marks)**

c) Schedules, serializability and isolation. **(5 marks)**

d) COMMIT, ROLLBACK, SAVEPOINT and staged or cascaded variants. **(5 marks)**

e) Locking levels, types and philosophies (optimistic versus pessimistic). **(5 marks)**

B5

Using your own simple examples and suitable diagrams, explain the following data modelling concepts.

a) Entity Relationship Diagrams (ERD), Entity Types, Entity Instantiations and Identifiers **(5 marks)**

b) Enhanced Entity Relationship Diagram (EERD), Super-Types and Sub-Types **(5 marks)**

c) Data Cubes & OLAP **(5 marks)**

d) Star Schemas and the role of normalization/de-normalization **(5 marks)**

e) Snowflake Schemas and the role of normalization/de-normalization **(5 marks)**